AD-A278 362

ION PAGE

Form Approved
OMB No. 0704-0188

...ge 1 hour per response, including the time for reviewing instructions, searching existing data sources, ... collection of information. Send comments regarding this burden estimate or any other aspect of this ... Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson ... nagement and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 22, 1994	annual technical 9/1/92--8/31/93

4. TITLE AND SUBTITLE
Neural models of motion perception

5. FUNDING NUMBERS
AFOSR F49620-92-J-0334
61103D
3484
S4

(2)

6. AUTHOR(S)
Stephen Grossberg and Ennio Mingolla, co-PIs

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Boston University
Department of Cognitive and Neural Systems
111 Cummington Street
Boston, MA 02215

8. PERFORMING ORGANIZATION REPORT NUMBER
AEOSR-TR- 94 0142

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Office of Scientific Research
Bolling AFB, DC 20332

Dr John F. Tangney

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-92-J-0334

11. SUPPLEMENTARY NOTES

DTIC
ELECTE
APR 21 1994
S F D

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Eight research projects supported by this grant during the reporting period have resulted in three refereed publications, one under review, one book chapter, and three conference papers. Areas of research included design and simulation of neural architectures for: (1) multichannel data fusion; (2) object recognition and image understanding; (3) development and refinement of algorithms for segmentation, boundary completion, and featural filling-in based on BCS/FCS architectures; (4) network design and simulations of an architecture for breaking of unwanted persistence (hysteresis) of visual segmentations; (5) design of a network architecture for explaining human capabilities for efficient detection of targets in clutter; (6) design and execution of human psychophysical experiments for constraining development of the BCS; (7) design and simulation of a network architecture for enhancing featural contrast and boundary localization at line-ends and corners through a novel circuit analog of V1 to lateral geniculate nucleus feedback; and (8) relation of hyperacuity and illusory contour data.

14. SUBJECT TERMS

DTIC QUALITY INSPECTED 3

15. NUMBER OF PAGES
11 pages

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

ANNUAL TECHNICAL REPORT

Contract AFOSR F49620-92-J-0334

NEURAL MODELS OF MOTION PERCEPTION

September 1, 1992—August 31, 1993

Co-Principal Investigators:
Stephen Grossberg and Ennio Mingolla
Boston University
Department of Cognitive and Neural Systems
111 Cummington Street
Boston, MA 02215
(617) 353-7857

ABSTRACT

Eight research projects supported by this grant during the reporting period have resulted in three refereed publications, one under review, one book chapter and three conference papers. Areas of reseach included design and simulation of neural architectures for: (1) multichannel data fusion; (2) object recognition and image understanding; (3) development and refinement of algorithms for segmentation, boundary completion, and featural filling-in based on BCS/FCS architectures; (4) network design and simulations of an architecture for breaking of unwanted persistence (hysteresis) of visual segmentations; (5) design of a network architecture for explaining human capabilities for efficient detection of targets in clutter; (6) design and execution of human psychophysical experiments for constraining development of BCS; (7) design and simulation of a network architecture for enhancing featural contrast and boundary localization at line-ends and corners through a novel circuit analog of V1 to lateral geniculate nucleus feedback; and (8) relation of hyperacuity and illusory contour data.

TABLE OF CONTENTS

PUBLICATIONS PARTIALLY SUPPORTED BY CONTRACT AFOSR F49620-92-J-0334

SEPTEMBER 1, 1992—AUGUST 31, 1993

ARTICLES

1. Asfour, Y.R., Carpenter, G.A., Grossberg, S., and Lesher, G.W. (1993). Fusion ARTMAP: A neural network architecture for multi-channel data fusion and classification. **Technical Report CAS/CNS-TR-93-006**, Boston University. In **Proceedings of the world congress on neural networks**, Portland, **II**, 210–215. Hillsdale, NJ: Erlbaum Associates. (%@+*)
2. Carpenter, G.A., Grossberg, S., and Lesher, G.W. (1993). The what-and-where filter: A spatial mapping neural network for object recognition and image understanding. **Technical Report CAS/CNS-TR-93-043**, Boston University. Submitted for publication. (%@+*)
3. Francis, G., Grossberg, S., and Mingolla, E. (1993). Cortical dynamics of feature binding and reset: Control of visual persistence. *Vision Research*, in press. (&%+*)
4. Francis, G., Grossberg, S., and Mingolla, E. (1993). Dynamic formation and reset of coherent visual segmentations by neural networks. To appear in **Artificial neural networks with applications in speech and vision**. London: Chapman and Hall, in press. (&%+*)
5. Gove, A., Grossberg, S., and Mingolla, E. (1993). Brightness perception, illusory contours, and corticogeniculate feedback. In **Proceedings of the world congress on neural networks**, Portland, **I**, 25–28. Hillsdale, NJ: Erlbaum Associates. (&%*)
6. Grossberg, S., Mingolla, E., and Ross, W.D. (1993). A neural theory of attentive visual search: Interactions of visual, spatial, and object representations. *Psychological Review*, in press. (&%+*)
7. Grossberg, S., Mingolla, E., and Williamson, J. (1993). Processing of synthetic and aperture radar images by a multiscale Boundary Contour System and Feature Contour System. In **Proceedings of the world congress on neural networks**, Portland, **III**, 785–788. Hillsdale, NJ: Erlbaum Associates. (&%*)
8. Lesher, G.W. and Mingolla, E. (1993). The role of edges and line-ends in illusory contour formation. *Vision Research*, **33**, 2253–2270. (&+*)

& Also supported in part by the AFOSR.

% Also supported in part by ARPA.

@ Also supported in part by British Petroleum. (nonrenewable)

+ Also supported in part by the National Science Foundation. (expired)

* Also supported in part by the Office of Naval Research.

RESEARCH SUMMARIES

1. Multichannel data fusion by a self-organizing network for recognition and prediction [Article 1]

Fusion ARTMAP is a self-organizing neural network architecture for multi-channel, or multi-sensor, data fusion. Single-channel Fusion ARTMAP is functionally equivalent to Fuzzy ART during unsupervised learning and to Fuzzy ARTMAP during supervised learning. The network has a symmetric organization such that each channel can be dynamically configured to serve as either a data input or a teaching input to the system. An ART module forms a compressed recognition code within each channel. These codes, in turn, become inputs to a single ART system that organizes the global recognition code. When a predictive error occurs, a process called parallel match tracking simultaneously raises vigilances in multiple ART modules until reset is triggered in one of them. Parallel match tracking hereby resets only that portion of the recognition code with the poorest match, or minimum predictive confidence. This internally controlled selective reset process is a type of credit assignment that creates a parsimoniously connected learned network. Fusion ARTMAP's multi-channel coding is illustrated by simulations of the Quadruped Mammal database.

2. Object recognition and image understanding [Article 2]

The What-and-Where filter forms part of a neural network architecture for spatial mapping, object recognition, and image understanding. The Where filter responds to an image figure that has been separated from its background. It generates a spatial map whose cell activations simultaneously represent the position, orientation, and size of the figure (where it is). This spatial map may be used to direct spatially localized attention to these image features. A multiscale array of oriented detectors, followed by competitive interactions between position, orientation, and size scales, is used to define the Where filter. The Where filter may be used to transform the image figure into an invariant representation that is insensitive to the figure's original position, orientation, and size. This invariant figural representation forms part of a system devoted to attentive object learning and recognition (what it is). The Where spatial map of all the figures in an image, taken together with the invariant recognition categories that identify these figures, can be used to learn multidimensional representations of objects and their spatial relationships for purposes of image understanding. The What-and-Where filter is inspired by neurobiological data showing that a Where processing stream in the cerebral cortex is used for attentive spatial localization and orientation, whereas a What processing stream is used for attentive object learning and recognition.

3. Processing of synthetic aperture radar images by a multiscale boundary segmentation and surface representation architecture [Article 7]

A multiscale image processing algorithm based on the Boundary Contour System (BCS) and Feature Contour System (FCS) neural network models of preattentive vision, developed at Boston University's Center for Adaptive Systems and Department of Cognitive and Neural Systems, has been transferred to MIT's Lincoln Laboratory and applied to large images containing range data gathered by a synthetic aperture radar (SAR) sensor. Researchers at

Lincoln Laboratory have in turn supplied enhanced versions of that software to clients at other laboratories. The goal of the algorithm is to make structures such as motor vehicles, roads, or buildings more salient and more interpretable to human observers than they are in the original imagery. Early automatic gain control by shunting center-surround networks compresses signal dynamic range while performing local contrast enhancement. Subsequent processing by filters sensitive to oriented contrast, including short-range competition and long-range cooperation, segments the image into regions. The segmentation is performed by three "copies" of the BCS and FCS, of small, medium, and large scales, wherein the "short-range" and "long-range" interactions within each scale occur over smaller or larger image distances, corresponding to the size of the early filters of each scale. Finally, a diffusive filling-in operation within the segmented regions generates surface representations of visible structures. The combination of BCS and FCS helps to locate and enhance structure over regions of many pixels, without the resulting blur characteristic of approaches based on low spatial frequency filtering alone.

4. Dynamic reset of boundary segmentations in response to rapidly changing imagery [Articles 3–4]

An analysis of the reset of visual cortical circuits responsible for the binding or segmentation of visual features into coherent visual forms yielded a model that explains properties of visual persistence described in Francis, Grossberg, and Mingolla (in press). The reset mechanisms prevent massive smearing of visual percepts in response to rapidly moving images. The model simulates relationships among psychophysical data showing inverse relations of persistence to flash luminance and duration, greater persistence of illusory contours than real contours, a U-shaped temporal function for persistence of illusory contours, a reduction of persistence due to adaptation with a stimulus of like orientation, an increase of persistence due to adaptation with a stimulus of perpendicular orientation, and an increase of persistence with spatial separation of a masking stimulus. The model suggests that a combination of habituative, opponent, and endstopping mechanisms prevent smearing and limit persistence.

The model consists of the BCS with habituative chemical transmitters embedded at the interface of its complex cells and hypercomplex cells. Thus *all* the properties used in image processing applications of the BCS are retained in the present model, which provides the additional advantage of rapidly resetting *only* those boundary groupings of a processed scene which are *changing* in a time-varying environment.

5. A network architecture to rapidly search and detect visual targets in clutter [Article 6]

Visual search data were given a unified quantitative explanation by a model of how spatial maps in the parietal cortex and object recognition categories in the inferotemporal cortex deploy attentional resources as they reciprocally interact with visual representations in the prestriate cortex, as described in Grossberg, Mingolla, and Ross (in press). The model visual representations are organized into multiple boundary and surface representations. Visual search in the model is initiated by organizing multiple items that lie within a given boundary or surface representation into a candidate search grouping. These items are compared with object recognition categories to test for matches or mismatches. Mismatches can trigger deeper searches and recursive selection of new groupings until a target object is

identified. This search model is algorithmically specified to quantitatively simulate search data using a single set of parameters, as well as to qualitatively explain a still larger data base, including data of Aks and Enns (1992), Bravo and Blake (1990), Chellazzi, Miller, Duncan, and Desimone (1993), Cohen and Ivry (1991), Egeth, Virzi, and Garbart (1984), Enns and Rensink (1990), He and Nakayama (1992), Humphreys, Quinlan, and Riddoch (1989), Mordkoff, Yantis, and Egeth (1990), Nakayama and Silverman (1986), Treisman and Gelade (1980), Treisman and Sato (1990), Wolfe, Cave, and Franzel (1989), and Wolfe and Friedman-Hill (1992). The model hereby provides an alternative to recent variations on the Feature Integration and Guided Search models, and grounds the analysis of visual search in neural models of preattentive vision, attentive object learning and categorization, and attentive spatial localization and orientation.

6. Human psychophysical experiments on boundary segmentation [Article 8]

Lesher and Mingolla (1993) showed that illusory contours can be induced along directions approximately collinear to edges or approximately perpendicular to the ends of lines. Using a rating scale procedure, they explored the relation between the two types of inducers by systematically varying the thickness of inducing elements to result in varying amounts of "edge-like" or "line-like" induction. Inducers for the illusory figures consisted of concentric rings with arcs missing. Observers judged the clarity and brightness of illusory figures as the number of arcs, their thicknesses, and spacing were parametrically varied. Degree of clarity and amount of induced brightness were both found to be inverted-U functions of the number of arcs. These results mandate that any valid model of illusory contour formation must account for interference effects between parallel lines or between those neural units *responsible for completion of boundary signals in directions perpendicular* to the ends of thin lines. Line width was found to have an effect on both clarity and brightness, a finding inconsistent with those models which employ *only* completion perpendicular to inducer orientation. Subsequent research reported in Lesher (1993) showed that the BCS could fit the data of the Lesher and Mingolla (1993) experiment.

7. A link between brightness perception, illusory contours, and binocular corticogeniculate feedback

As reported in Gove, Grossberg, and Mingolla (1994), many illusory contour displays induce apparent brightness along the ends of thin lines. "Brightness buttons" are usually described as unnoticed for sinlge lines, but effective in producing the enhanced brightness inside the illusory contours induced by Ehrenstein patterns. No satisfactory neural mechanism for brightness buttons has yet been suggested. We propose that they are consequences of corticogeniculate feedback whose primary functional role is to selectively prime monocular LGN cells whose activation is consistent with fused binocular activation of cortical V1 cells. We simulated a model of neural circuitry of LGN and V1. Model LGN relay cells receive input from retinal cells, positive feedback from oriented V1 cells, and negative feedback from LGN interneurons, which also receive cortical feedback. Brightness button signals can be generated in two ways consistent with reported physiology: (1) Excitatory feedback from cortical end-stopped cells can enhance LGN cell activity near line ends; (2) Net inhibitory feedback from long-field cells, modulated by LGN interneurons, can suppress activity in LGN cells coding the sides of lines, making brightness contrast at line ends relatively stronger. A combination of the two mechanisms has the same properties. Our research shows that

brightness enhancement of illusory figures that are induced at line ends may reflect corticogeniculate feedback mechanisms. These mechanisms select monocular LGN cells whose activation is consistent with that of the binocular cortical cells that are used to form the illusory contours.

8. Relation of hyperacuity and illusory contour data

Lesher's (1993) dissertation contains (among other projects) simulations describing how the BCS can fit the illusory contour data of Project 6 in a manner that unifies the treatment of hyperacuity data and illusory contour formation, as first described by Grossberg (1987). Tradeoffs in network design for optimal spatial resolution and for reconciling long-range contextual information with local data are thereby accorded a unified treatment.

References

Aks, D.J. and Enns, J.T. (1992). Visual search for direction of shading is influenced by apparent depth. *Perception and Psychophysics*, **52**, 63–74.

Bravo M. and Blake R. (1990). Preattentive vision and perceptual groups. *Perception*, **19**, 515–522.

Chelazzi, L., Miller, E.K., Duncan, J., and Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex. *Nature*, **363**, 345–347.

Cohen, A. and Ivry, R.B. (1991). Density effects in conjunctive search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 891–901.

Egeth, H., Virzi, R.A., and Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 32–39.

Enns, J.T. and Rensink, R.A. (1990). Influence of scene-based properties on visual search. *Science*, **247**, 721–723.

Gove, A., Grossberg, S., and Mingolla, E. (1994). A link between brightness perception, illusory contours, and binocular corticogeniculate feedback. *Association for Research in Vision and Ophthalmology Abstracts*, in press.

Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception, Parts I and II. *Perception and Psychophysics*, **41**, 87–158.

He, Z.J. and Nakayama, K. (1992). Surface features in visual search. *Nature*, **359**, 231–233.

Humphreys, G.W., Quinlan, P.T., and Riddoch, M.J. (1989). Grouping processes in visual search: Effects with single and combined feature targets. *Journal of Experimental Psychology: General*, **118**, 258–279.

Lesher, G.W. (1993). Neural networks for vision and pattern recognition: Boundary completion, spatial mapping, and multidimensional data fusion. PhD Thesis, Boston University. Ann Arbor: University Microfilms 9401999.

Mordkoff, J.T., Yantis, S., and Egeth H.E. (1990). Detecting conjunctions of color and form in parallel. *Perception and Psychophysics*, **5**, 157–168.

Nakayama, K. and Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, **320**, 264–265.

Treisman, A. and Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, **16**, 97–136.

Treisman, A. and Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 459–478.

Wolfe, J.M., Cave, K.R., and Franzel, S.L. (1989). Guided search: An alternative to the feature integration model of visual search. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 419–433.

Wolfe, J.M. and Friedman-Hill, S.R. (1992). Part-whole relationships in visual search. *Investigative Opthalmology and Visual Science*, **33**, 1355.

STUDENT SUPPORT

Gregory Lesher and Jeff Yuan, graduate students in the Department of Cognitive and Neural Systems, received support from this grant.

Lesher received his PhD degree in Cognitive and Neural Systems in May, 1993. The topic of his dissertation was "Neural networks for vision and pattern recognition: Boundary completion, spatial mapping, and multidimensional data fusion." He is currently working as a consultant for the Pacific Sierra Research Corporation.

Yuan continues to work towards the completion of his PhD degree.

Approved for public release;
distribution unlimited.

AIR FORCE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12
distribution in unlimited.
Joan Boggs
STINFO Program Manager